Integrating genomics in population models to forecast translocation success

Travis Seaborn\textsuperscript{1,2}, Kimberly R. Andrews\textsuperscript{3}, Cara V. Applestein\textsuperscript{4}, Tyler M. Breech\textsuperscript{5}, Molly J. Garrett\textsuperscript{1}, Andrii Zaiats\textsuperscript{4}, T. Trevor Caughlin\textsuperscript{4}

Whole-genome sequencing is revolutionizing our understanding of organismal biology, including adaptations likely to influence demographic performance in different environments. Excitement over the potential of genomics to inform population dynamics has prompted multiple conservation applications, including genomics-based decision-making for translocation efforts. Despite interest in applying genomics to improve translocations, there is a critical research gap: we lack an understanding of how genomic differences translate into population dynamics in the real world. We review how genomics and genetics data could be used to inform organismal performance, including examples of how adaptive and neutral loci have been quantified in a translocation context, and future applications. Next, we discuss three main drivers of population dynamics: demographic structure, spatial barriers to movement, and introgression, and their consequences for translocations informed by genomic data. Finally, we provide a practical guide to different types of models, including size-structured and spatial models, that could be modified to include genomics data. We then propose a framework to improve translocation success by repeatedly developing, selecting, and validating forecasting models. By integrating lab-based and field-collected data with model-driven research, our iterative framework could address long-standing challenges in restoration ecology, such as when selecting locally adapted genotypes will aid translocation of plants and animals.

Key words: assisted gene flow, assisted migration, conservation genetics, conservation genomics, iterative modeling, reintroduction

Implications for Practice

- Genomic data will continue to become available for species of conservation concern, including species targeted for translocation efforts.
- While genomic data can identify local adaptations relevant for spatially targeted translocation, ecological complexity, including demographic structure and landscape context, could undermine success of efforts based primarily on genomics.
- Data on translocation outcomes, including population size and genetic diversity, could be used to assess the predictive power of gene-by-environment interactions for translocation success.
- Forecasting models that iterate between lab-based predictions and field-measured outcomes will improve our capacity to plan translocations using genomics data, but we highlight continued challenges for these methods to be widely applied.

Introduction

Translocation of plants and animals is central to efforts to restore biodiversity. Species of conservation concern often face inbreeding depression, reduced adaptive capacity, local extinctions, and other consequences of small population sizes. However, despite decades of effort and enormous amounts of funding spent on translocation efforts to combat these concerns, success, characterized by establishing a self-sustaining, free-ranging, viable population (Griffith et al. 1989; Fischer & Lindenmayer 2000), is notoriously unpredictable, and examples of failure are common (Godefroid et al. 2011; Drayton & Primack 2012; Cochran-Biederman et al. 2015). This lack of consistent success has prompted efforts to better forecast outcomes of translocations. Accurate forecasting ability is even more critical as we enter an era of unprecedented climate change; predicting species’ responses to shifting environmental conditions is necessary for matching locally adapted populations to similar future conditions. As genome-wide sequence data becomes increasingly available for non-model organisms...
due to rapid advances in DNA sequencing technologies, genomic data could revolutionize translocation predictions and planning. Whole genome sequencing has the potential to enable identification of local adaptations and to increase the number of neutral markers available for population genetic analyses. However, while predicting future translocation success is a primary goal of genome-wide sequencing for conservation, frameworks to quantify how genomic differences impact population growth rates of wild plants and animals remain scarce.

Translocation forecasts informed by genetic information build off a successful track record of using population genetic analyses to address questions in conservation biology either before or after translocation (Boxes 1, 2). Genetic metrics used to aid in selecting source populations, and post-translocation monitoring, include genetic diversity, population structure, inbreeding coefficients, and effective population size. For instance, source populations for translocations are often selected based on high levels of genetic diversity and moderate levels of genetic similarity to the recipient population (IUCN/SSC 2013; Houde et al. 2015). Genetic data collected prior to translocations can be compared with post-translocation genetic data to monitor post-translocation outcomes (e.g. Madsen et al. 1999; McGlaughlin et al. 2002; De Barba et al. 2010). However, post-translocation genetic monitoring is rare (Weeks et al. 2011; Ewen et al. 2014) despite its strong potential benefits.

**Box 1** Pre-translocation data from common garden experiments improve seed sourcing during sagebrush restoration

![Sagebrush plants from source populations throughout the Great Basin planted in a common garden experiment in Orchard, ID. Aerial photo courtesy of Donna Delparte.](image)

Big sagebrush (*Artemisia tridentata*) in the Great Basin provides an example of how pre-translocation data is being applied to improve translocation success. A large-scale analysis of sagebrush reseeding treatments in post-fire rehabilitation projects revealed that the majority of seeding treatments failed to increase cover or density of sagebrush (Knutson et al. 2014). One potential reason for this large-scale failure is inappropriate seed sourcing (selecting from populations which were not genetically adapted to the microclimates in which they were seeded). Follow-up common garden studies have shown that climate characteristics of source population habitats can predict genetic compatibility with new locations (Chaney et al. 2017). While understanding genomic adaptation to specific climates in sagebrush can improve seed sourcing, post-translocation monitoring will continue to play a role in assessing translocation success beyond the common garden. Iterative forecasting models which simulate individuals could quantify the demographic benefits of local adaptation for translocation in changing landscapes of the American West.

**Box 2** Post-translocation monitoring to detect introgression between native and introduced trout species

![Westslope Cutthroat Trout from the St. Joe River, Idaho, USA. Translocations of native trout species are ongoing to restore genetic integrity after global introductions of Coastal Rainbow Trout (*Oncorhynchus mykiss irideus*). During Westslope Cutthroat Trout (*Oncorhynchus clarkii lewisi*) translocation in the American West, post-translocation genomic data provide critical assessments of potential introgression. Fisheries scientists use genomic data to quantify relationships between trout stocking and hybridization, such as genome-wide selection against introgression or introgression even after translocations have been ended (Muhlfeld et al. 2015; Kovach et al. 2016). By measuring hybridization rates at multiple time points, researchers could use iterative forecasting models to predict the levels of hybridization and outbreeding depression for future translocation (or eradication) scenarios.

These forecasting models could be created by using demo-genetic, individual-based models. Although this has been done in the past with limited numbers of markers with modeling frameworks like CDMetaPOP (see Nathan et al. 2019), current expansions of modeling frameworks will allow a greater number of loci to be used to increase statistical power for quantifying introgression, although pinpointing loci directly involved in inbreeding depression may still be challenging (Kardos et al. 2016). This expansion to modeling frameworks highlights the computational biology needs to advance the incorporation of genomic data in translocation forecasts.
Whereas genetic analyses have traditionally used small numbers of genetic markers, rapid advancements in DNA sequencing technologies are dramatically increasing the accessibility of data from thousands of genetic markers across the genome (Allendorf et al. 2010; Andrews et al. 2016; Hohenlohe et al. 2018). Using a large number of genetic markers substantially increases the statistical power for population genetic analyses. In addition, genome-wide sequence data allow the identification of genomic variants responsible for adaptation to local habitats (Savolainen et al. 2013), and this information can help predict whether individuals or populations will succeed if translocated to a new habitat. Importantly, genomic data can distinguish adaptations with a genomic basis from those resulting from phenotypic plasticity; adaptations with a genomic basis are more likely to be predictive of long-term translocation success (Aitken & Whitlock 2013).

We propose that existing models for biological complexity in population dynamics of translocated plants and animals could serve as an integrative framework for applying genomic data to translocations. This review assesses the integration between translocations and genomic data. We do this by (1) discussing benefits for using genomic data with translocation events, both before and after the event (“pre-translocation” and “post-translocation”); (2) reviewing why translocations may fail when only genomic data are incorporated in forecasting efforts; and (3) arguing for the increased use of forecasting models that incorporate genomic data to inform translocation efforts. We use the broad term of “translocation” to encompass the concepts underlying the wide range of translocation-related terms currently in use (Table 1).

### Table 1. Definitions of translocation-related terms.

<table>
<thead>
<tr>
<th>Translocation Term</th>
<th>Definition</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assisted gene flow</td>
<td>Managed movement of individuals into populations to reduce local maladaptation to climate of other environmental changes</td>
<td>Whiteley et al. (2015)</td>
</tr>
<tr>
<td></td>
<td>Human-mediated gene flow</td>
<td>Allendorf et al. (2012)</td>
</tr>
<tr>
<td></td>
<td>The managed translocation of individuals within the current species range to facilitate rapid adaptation to climate change</td>
<td>Aitken and Ste-Marie et al. (2011)</td>
</tr>
<tr>
<td>Assisted migration</td>
<td>The intentional movement of species or populations, generally within their current range, in response to observed or anticipated climate change</td>
<td>Ste-Marie et al. (2011)</td>
</tr>
<tr>
<td></td>
<td>Movement from historical range to new areas based on climate shift</td>
<td></td>
</tr>
<tr>
<td>Genetic rescue</td>
<td>An increase in population fitness (growth) owing to immigration of new alleles</td>
<td>Whiteley et al. (2015)</td>
</tr>
<tr>
<td></td>
<td>The recovery in the average fitness of individuals through gene flow into small populations, typically following a fitness reduction due to inbreeding depression</td>
<td>Allendorf et al. (2012)</td>
</tr>
<tr>
<td>Managed relocation</td>
<td>An intervention technique aimed at reducing negative effects of climate change on defined biological units such as populations, species, or ecosystems. It involves the intentional movement of biological units from current areas of occupancy to locations where the probability of future persistence is predicted to be higher</td>
<td>Richardson et al. (2009)</td>
</tr>
<tr>
<td>Conservation translocation</td>
<td>The movement of individuals from one population (or location) to another that is usually intended to achieve either genetic or demographic rescue of an isolated population, or to allow adaptation to a rapidly changing climate</td>
<td>Allendorf et al. (2012)</td>
</tr>
<tr>
<td></td>
<td>The intentional movement of living organisms from one area to another</td>
<td>Seddon et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>A general term of the transfer by human agency of any organism(s) from one place to another (based on NCC 1990)</td>
<td>Hodder and Bullock (1997)</td>
</tr>
<tr>
<td></td>
<td>The intentional release of animals to the wild in an attempt to establish, reestablish, or augment a population and may consist of more than one release</td>
<td>Griffith et al. (1989)</td>
</tr>
<tr>
<td></td>
<td>Deliberate movement of organisms from one site for release in another, the human-mediated movement of living organisms from one area, with release in another; the intentional movement and release of a living organism where the primary objective is a conservation benefit: this will usually comprise improving the conservation status of the focal species locally or globally, and/or restoring natural ecosystem functions or processes</td>
<td>IUCN/SSC (2013)</td>
</tr>
<tr>
<td>Conservation reintroduction</td>
<td>The introduction of a species or population into a historic habitat from which it had been previously extirpated</td>
<td>Allendorf et al. (2012)</td>
</tr>
<tr>
<td></td>
<td>The intentional release of a species into its historical range where it has become extirpated. Reintroduction is distinct from reinforcement, which involves translocation of organisms to existing populations of the same species</td>
<td>IUCN/SSC (2013); He et al. (2016)</td>
</tr>
<tr>
<td></td>
<td>The intentional movement and release of an organism inside its indigenous range from which it has disappeared</td>
<td>IUCN/SSC (2013)</td>
</tr>
</tbody>
</table>

### Review of Genomic Data in Translocation Efforts

#### Pre-Translocation Data

**Neutral Loci.** Neutral loci, which are genetic loci that do not directly influence fitness, have informed translocation efforts in multiple ways. First, neutral loci can delineate populations to generate an understanding of population structure and connectivity (Schwartz 2005; Greenwald 2010). Second, neutral
Adaptive Loci. In contrast to neutral loci, adaptive loci represent areas of the genome directly related to fitness, including reproduction and survival. Several approaches have been developed for using DNA sequence data to identify genomic adaptation within individuals and populations, which can assist in translocation decision-making. These approaches focus on identifying genomic regions associated with phenotypes or environmental variables. Methods such as quantitative trait locus (QTL) mapping (Stinchcombe & Hoekstra 2008) and genome-wide association studies (GWAS; Korte & Farlow 2013) are considered top-down approaches that can identify regions of the genome related to local adaptation by identifying associations between genotypes and phenotypes. In contrast, $F_{ST}$ outlier tests identify locally adaptive genetic variants under divergent adaptive pressures by identifying genetic regions that differ in allele frequencies between populations (Francois et al. 2016). This contrasts with environmental associational analyses (EAAs), which identify genetic variants strongly correlated with specific environmental conditions using methods parallel to those for identifying genetic variants associated with phenotypic traits in QTL analyses (Rellstab et al. 2016). These methods help identify loci directly to traits, which often may not be possible for other methods which identify loci but do not determine the translation of genotypes to phenotypes. Methods to identify loci associated with environmental conditions can be either multivariate, such as redundancy analysis (Forster et al. 2018), or univariate, such as latent factor mixed models (Frichot et al. 2013), and many EAA methods explicitly control for neutral population genetic structure (Rellstab et al. 2016). Often multiple methods are used for identifying adaptive loci, and researchers focus downstream analyses on loci that are found to be significant across multiple methods (Dorant et al. 2020; Massatti & Knowles 2020).

Incorporation of adaptive genomic variation into management decision-making is a relatively new endeavor, and best practices remain the subject of discussion and debate (Funk et al. 2019; Hoelzel et al. 2019). For example, genomic signatures of adaptation can be difficult to distinguish from genomic signatures of drift or purifying selection, or could have resulted from historic environmental conditions that have since disappeared; in these cases, loci exhibiting genomic signatures of adaptation may not be useful for management decision-making (Kardos & Shafer 2018). Nonetheless, molecular and analytical tools for identifying and understanding adaptive genomic variation are rapidly advancing, and incorporation of adaptive genomic information is increasingly being recognized as a valuable tool for translocations and other management applications (Funk et al. 2019). For example, Shryock et al. (2017) used a combination of EAAs and $F_{ST}$ outlier tests to identify genetic loci associated with climate covariates for two desert shrub species of high restoration value. This information was used to develop models for defining the areas where seeds are collected for translocation (“seed source zones”) to facilitate seed collection. Similar genomic approaches have identified genetic loci associated with climate covariates in tree species of conservation interest (Steane et al. 2014; Martins et al. 2018). Whole-genome sequences have also identified genetic loci associated with local adaptation in animal species, including pinpointing the most adapted individuals of Gunnison and Greater Sage-Grouse populations for spatially targeted translocation (Oh et al. 2019).

Post-Translocation Data

Neutral Loci. Comparisons of genetic diversity before and after translocations can be used to evaluate translocation success. Applications include assessing the short-term success of translocations to supplement declining and inbred populations by quantifying whether translocation increased genetic diversity (Madsen et al. 1999). Comparing neutral markers between historical and contemporary populations can reveal long-term costs and benefits of translocation efforts. For example, neutral loci were used to demonstrate that admixture between wild populations and translocated hatchery fish has likely resulted in declining genetic diversity in brown trout (Salmo trutta) populations (Hansen et al. 2009). Post-translocation monitoring of genetic diversity can also inform the design of future translocation efforts by establishing thresholds for the minimum population size needed to maintain genetic diversity (McGlaughlin et al. 2002). Analysis of genome-wide sequence data will enable rapid identification of declining genetic diversity in threatened populations (i.e. genetic erosion). New metrics are in development to quantify genetic erosion, such as analyses of runs of homozygosity (ROH; Leroy et al. 2018). ROH can potentially reveal early signs of inbreeding within populations with higher precision than traditional microsatellite markers (Leroy et al. 2018). Overall, post-translocation analyses can inform restoration practitioners about the genetic state of translocated populations with concrete recommendations and lessons for future interventions (Fant et al. 2013).

Pedigree assignments based on genetic data can also be beneficial for monitoring newly established populations. For example, multi-year post-translocation monitoring of a reintroduced brown bear population in the Italian Alps revealed that one dominant male had sired all cubs born within the last 4 years, leading to increased relatedness and decreased genetic diversity in the population (De Barba et al. 2010). Without genetic monitoring, this reintroduction effort would have been deemed a success based solely on the population’s increased size. Pedigree analyses based on genetic data have also been used to monitor
reintroduction efforts for gray wolves (Stenglein et al. 2011) and Columbia Basin pygmy rabbits (DeMay et al. 2017). Pedigree analysis methods also have promise for selecting which individuals to breed to minimize relatedness (Galla et al. 2020).

Adaptive Loci. Adaptive loci have been used much less frequently to monitor post-translocation events than in the pre-translocation planning process (Table 2). These markers may allow future monitoring to look for adaptive divergence compared to the source populations. In addition, adaptive loci could be used to monitor adaptive changes within the new population through time.

What Factors Could Undermine Forecasts Informed by Genomic Data?

Demographic Structure

Genomic predictions that do not take demographic structure into account could have poor predictive ability. Populations of plants and animals typically consist of individuals at varying stages of the life cycle, including different sizes, stages, and ages. Conservation biologists have long recognized that demographic structure can interact with population dynamics with significant consequences for management interventions (Caswell 2000). For similar reasons, reintroduction efforts that focus on translocating small individuals (e.g. seedlings or fish fry) may appear to fail for up to several decades while larger individuals establish (Shriver et al. 2019). The age distribution of translocated individuals also has a strong impact on population growth (Linklater et al. 2011; Shier & Swaisgood 2012), as does the sex of translocated individuals, since sex ratio and breeding systems interact with population dynamics (Bosé et al. 2007). Signals of local adaptation may also be more strongly expressed in traits at certain life history stages, with consequences for population dynamics in heterogeneous environments (Rice & Knapp 2008; Rees & Ellner 2016). The importance of demographic structure for population dynamics points to the need to consider genomic differences in the context of organism life cycles.

Barriers to Movement

Genomic-informed predictions that do not account for landscape context and configuration could also have poor translocation success rates. Geographic barriers to movement are ubiquitous, and translocation decisions that assume individuals will disperse freely could have unexpected outcomes. Human-made structures (e.g. highways; De Barba et al. 2010) can act as connectivity barriers for reintroduced populations. Natural landscape features can also limit gene flow in reintroduced populations; for example, microsatellite analyses revealed that historical translocation of fishers (Pekanina pennanti), a carnivorous mustelid, in the Northeast was impeded by lakes

<table>
<thead>
<tr>
<th>Biological Process</th>
<th>Type of Model</th>
<th>Genomic Data Inputs</th>
<th>Additional Data Inputs</th>
<th>Pre- or Post-Translocation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local adaptation and gene flow</td>
<td>Dispersal</td>
<td>Simulated adaptive loci</td>
<td>Biophysical environmental characteristics</td>
<td>Pre</td>
<td>Quigley et al. (2019); Dudas et al. (2017)</td>
</tr>
<tr>
<td>Local adaptation and complex demographics</td>
<td>Logistic growth model</td>
<td>Adaptive loci</td>
<td>Demographics, temperature, artificial migration</td>
<td>Pre</td>
<td>Bay et al. (2017)</td>
</tr>
<tr>
<td>Gene flow</td>
<td>Dispersal</td>
<td>Simulated neutral chromosome</td>
<td>Abundance, habitat suitability</td>
<td>Post</td>
<td>Rougemont et al. (2019); Malone et al. (2018)</td>
</tr>
<tr>
<td>Local adaptation</td>
<td>Source suitability decision framework</td>
<td>Neutral loci</td>
<td></td>
<td>Pre</td>
<td>Hogg (2020)</td>
</tr>
<tr>
<td>Local adaptation and complex demographics</td>
<td>Individual-based model</td>
<td>Adaptive and neutral loci</td>
<td>Demographics, habitat</td>
<td>Either</td>
<td></td>
</tr>
<tr>
<td>Local adaptation</td>
<td>Species distribution model</td>
<td>Relate locally adapted genotypes to environmental variation</td>
<td>Abundance or occurrence of species in environment</td>
<td>Pre</td>
<td>Jonsson et al. (2018)</td>
</tr>
<tr>
<td>Complex demographics</td>
<td>Gradient forest regression</td>
<td>Adaptive loci</td>
<td>Environmental associations</td>
<td>Pre</td>
<td>Bay et al. (2018)</td>
</tr>
<tr>
<td>Complex demographics</td>
<td>Stage or size-structured population model</td>
<td>Genetic polymorphism</td>
<td>Data on stage-specific vital rates</td>
<td>Either</td>
<td>de Vries and Caswell (2019)</td>
</tr>
<tr>
<td>Complex demographics</td>
<td>Integrated population model</td>
<td>Could incorporate individual-level variables</td>
<td>Flexible; can incorporate multiple types of data</td>
<td>Post</td>
<td>Plard et al. (2019)</td>
</tr>
</tbody>
</table>
In addition, translocated individuals may move and respond to potential barriers differently than non-translocated individuals (Jesmer et al. 2018). In some cases, barriers that prevent gene flow can benefit translocation if barriers prevent introgression, the breeding of related taxa through hybridization (Harig et al. 2000).

Introgression

Introgression is another factor that could impede translocation success and planning by leading to rapid evolutionary changes, including genetic and phenotypic feedbacks (Brasier 2001; Dierking et al. 2014). For example, two species of Indian paintbrush, *Castilleja levisecta*, a federally threatened species, and *Castilleja hispida*, the host plant for a federally endangered butterfly, *Euphydryas editha taylori*, were accidentally planted in close proximity on prairie restoration sites in western Washington, leading to concerns over hybridization and ultimately the removal of some translocated individuals (Dunwiddie & Martin 2016). While introgression can undermine efforts to restore native populations, introgression can also have negative impacts on fitness (e.g. Muhlfeld et al. 2009; but see Frankham 2015). The likelihood of introgression, and its associated eco-evolutionary feedbacks, can be unpredictable, necessitating data-driven approaches to quantify hybridization.

Iterative Forecasting Models to Improve Translocations

Demographic structure, barriers to movement, introgression, and a myriad of other complications are likely to undermine forecasts of reintroduction success based primarily on genomic data. These factors raise the question of how to develop predictions of reintroduction success that integrate genomic data yet acknowledge ecological complexity. In other areas of science, model-based forecasts of future events play a major role in planning and decision-making (e.g. weather forecasts). Ecological forecasts are also under development, aided by recent efforts to evaluate ecological models in the context of how well models can extrapolate beyond the range of data used for model-fitting (Dietze et al. 2018). Integrating multiple sources of data, such as the potential performance of organisms under a range of environmental conditions and time series of population abundance, will play a key role in developing forecasting models. For example, adaptive genomic-based predictions of heat tolerance could be used to develop a model to predict organissmal abundance, and then independent field data on abundance could provide a test of model predictions (Fig. 1). With the increasing availability of existing genetic, environmental, and other data, along with increasing availability of novel genomic data, we are well-poised to develop forecasting models that iterate between model development and model testing (Draper et al. 2019).

Modeling Demographic Structure

Demographic models for conservation biology enable variation in stage, size, and age structure with potential to interact with local adaptation (Caswell 2000). Models that include demographic structure are necessary for long-term forecasts (Caughin et al. 2014, 2019) and can also produce more detailed recommendations for management, such as targeting particular demographic classes for management action. Incorporating genotypic variation as a source of phenotypic variation into population demographic models will lead to novel insights into eco-evolutionary dynamics with application to reintroduction (Metcalf & Pavard 2007). For example, demographic models that incorporated bodyweight (a metric of demographic structure) and a gene with two alleles that determine coat color were used to forecast the response of reintroduced wolf populations to environmental change (Coulson et al. 2011). Demographic structure is likely to be most important when there is an interaction between genotype and individual stage, size, sex, or age, in determining demographic rates, or when one stage is a demographic bottleneck (Easterling et al. 2000). Incorporating

![Figure 1. Conceptual diagram of genomic integration with iterative translocation forecasting models. Data below provide one example of several types of relevant data.](image)
measurements of demographic structure in pre-translocation data collection (e.g. measuring plant size in greenhouse experiments to determine the relationship between plant size and adaptation to stress) and post-translocation data on organismal abundance will aid our ability to develop forecasting models that include demographic structure.

Modeling Barriers to Movement

Spatially explicit models can further advance translocation planning and monitoring. Ensuring the translocated populations maintain connectivity with remnant populations is a strong foundation for successful restoration projects (Aavik & Helm 2018). Dispersal is a fundamental component of genetic connectivity that can readily be included in spatially explicit modeling and forecasting frameworks (Lethbridge & Strauss 2015; Landguth et al. 2017). Furthermore, incorporating genetic structure and gene flow into these frameworks, along with individual dispersal, can help improve the predictions of long-term population viability (Godefroid et al. 2011). Several modeling frameworks, including mechanistic dispersal kernels and spatiotemporal models, can account for spatial and temporal processes post-translocation. For example, spatial reaction–diffusion models (e.g. Wright-Fisher equation), originally developed to model the spread of new alleles in a population, can represent both genetic and demographic diffusion (Roques et al. 2016; Helley et al. 2017; Bergman et al. 2018). Modeling advection processes, including animal movement, can account for habitat heterogeneity in recovering landscapes (Moorcroft & Barnett 2008). When confronted with pre- and post-translocation genomic and spatial population data collected over time, such models represent a rigorous framework for forecasting population persistence under novel biotic and abiotic regimes.

Modeling Introgression

Modeling introgression with genomic data in tandem will also increase the success of translocation efforts. Introgression can have a wide array of impacts on survival and reproductive success (Baack & Rieseberg 2007; Colella et al. 2019) and is often considered in translocation planning (González-Trujillo et al. 2012; Meek et al. 2014). Genomic data provides increased power for detecting and quantifying introgression, and also enables assessment of whether selective pressures cause certain parts of the genome to introgress more or less than others (Rosenzweig et al. 2016; Martin & Jiggins 2017). This increase in power has allowed previously unknown introgression to be discovered (Twyford & Ennos 2012; Liu et al. 2015). In other cases, introgression levels have been lower than previously estimated (Hohenlohe et al. 2013). As more information regarding rates and prevalence of introgression become available, there will be increasing opportunities for modeling population-level consequences of introgression. Although evaluating introgression with translocations may be common, it is less common to use modeling to predict outcomes of translocations regarding introgression (but see Colella et al. 2019; Jager 2006). Depending on the translocation project’s goals, maintaining desired introgression levels would benefit from an iterative modeling approach, including pre-translocation predictions and post-translocation data on temporal samples of introgression occurrence. Eco-evolutionary models could also represent population

Figure 2. Example of pre- and post-translocation workflow integrating genomics into population models to maximize success. Thirty populations of a fish species from across an environmental gradient are in consideration for a translocation event to a site where the population of that species has been extirpated. Managers aim to choose fish for translocation that will lead to the establishment of a self-sustaining, viable population.

1. Pre-translocation empirical data collection: fin clips from all 30 populations for genomic analyses, water quality metrics and other site characteristics for environmental analyses such as reproductive rates.
2. Genomics benchwork: DNA extraction and sequencing
3. Empirical data processing: genotyping, followed by estimation of genetic diversity and Environmental Association Analysis to identify “adaptive” loci associated with relevant environmental parameters that vary across the populations and the translocation site.
4. Modelling framework to simulate translocation: track genetic diversity and adaptive loci allele frequencies in CDMetaPop or HexSim using simulations that vary parameters including the number, demographic composition, genomic composition, and geographic origin of translocated individuals, as well as the seasonal timing of the translocation event and the strength of selection on the adaptive loci.
5. Translocation event: Select individuals for the translocation based on outcomes of modeling scenarios.
6. Post-translocation empirical data collection: fin clips and age-structured population size estimate at translocation site
7. Genomics benchwork: DNA extraction and sequencing
8. Post-translocation data processing: genotyping, followed by estimation of genetic diversity and frequency of adaptive alleles
9. Updating models: forecasting future status, estimate model parameters that may need to be updated. Identify best parameters for another translocation event
10. Repeat steps 5-9 in an iterative approach until translocated population is self-sustaining and viable.
Creating Forecasting Models With Genomics: Two Approaches

We present two approaches for integrating genomic data into forecasting models. The first approach involves a two-step process, which begins with conducting genomic analyses to determine inputs for population models (Fig. 2). Individual-based models provide an example of a population model capable of simulating the genomic composition of multiple loci and individuals. However, individual-based models do not allow for the large number of putatively adaptive loci that often emerge from empirical genomic work, and instead simulate only a few adaptive loci (Table 2). In contrast, the incorporation of many neutral loci is more common across modeling frameworks. There are some exceptions; CDMetaPOP was recently expanded to allow for multilocus selection using a linear additive selection model for genotype-environment interactions which can account for many adaptive loci in addition to neutral loci (Landguth et al. 2020). HexSim can handle any number of neutral loci (Schumaker & Brookes 2018). RangeShifter has also recently had its modules expanded to handle pleiotropy and more complex genomic data (Bocedi et al. 2020). Nevertheless, selecting appropriate loci from full genome sequencing for simulation models remains a challenge. One potential solution is to reduce the dimensionality of genomic data using indices derived from multivariate ordination techniques, such as PCA (e.g. Milesi et al. 2019).

A second approach for integrating genomics into population models could involve jointly estimating regions of the genome important for near-term population dynamics and individual or population-level demography. Demographic models that incorporate functional linear models or universal differential equations can accommodate a large number of predictors (Rackauckas et al. 2020). For example, Teller et al. (2016) demonstrated how a suite of climate observations can be integrated in time into a function-valued covariate to explain plant growth. Analytically tractable population models that are generated from regression-type statistical models (e.g. integral projection models; Rees & Ellner 2016) present one opportunity to incorporate genomic information, estimated from functional linear models or random forest algorithm, into population growth rate. While this type of model has not yet been implemented, single-step approaches to fit genomics-informed population models will benefit from the rapidly developing field of multidimensional data analysis.

Model Complexity and Future Prospects

An appropriate model type for genomics-informed translocation and forecasting remains an open question. The answer will likely depend on the attributes of the ecological system and management goals, with pre- and post-translocation data being a key component of the conservation strategy. To begin the process, we recommend collaboration between ecological modelers and conservation practitioners to work through the process of developing the models. A wide range of models could forecast translocation outcomes (e.g. genetic diversity or abundance) from simple reaction–diffusion equations to complicated individual-based models that represent heterogeneity across scales (Table 2). Model choice is guided in part by available data. For example, a complicated individual-based model may be less appropriate if most parameters are unavailable (e.g. dispersal rates and connectivity). At a minimum, forecasting models for translocation should include the predicted abundance of a species post-translocation. Other key details may include spatial patterns in abundance and the relative frequency of different genotypes. Demographic mechanisms, including dispersal and age- or size-structured vital rates (growth, survival, and reproduction), are central to modeling population dynamics of many plants and animals. Some of these model types (e.g. matrix population models) have a long history in conservation biology but have not yet been used as a testing ground for the impacts of genomic differences. Other model types have played an important role in understanding eco-evolutionary dynamics in simplified conditions (e.g. individual-based models), but may present computational challenges when forecasting translocation outcomes over large geographic areas (Romero-Mujalli et al. 2019).

While models present a valuable tool to forecast the success of translocation efforts, there are hard limits to the complexity that can be included in models. More complicated modeling approaches face limits related to data availability and computing processing power (e.g. Christin et al. 2019). One approach is to create multiple models, potentially of different types, to represent different hypotheses and predictions to be evaluated. Model selection could then enable evaluation of this suite of models. The most rigorous test for model performance involves out-of-sample testing on different sites, time periods, or species (Roberts et al. 2017). The need for out-of-sample model prediction provides a ready avenue for integrating data collected before and after a translocation event (Boxes 1, 2). A transdisciplinary collaboration that brings together data streams from mechanistic experiments and long-term observational data, including agency partners, will play a key role in the iterative development of forecasting models.

One critical step after creating and evaluating the models is to take an iterative approach of adding new data and re-evaluating models through time (Fig. 1). This approach can take multiple forms and will help improve predictions. First, by collecting new data after the translocation event, parameters for the species in that location can be refined. For example, due to unforeseen relationships with the environment, it is possible that the original demographic parameters estimated from the source population, such as survivorship, may not accurately reflect the new location. Second, besides refining with new data measured at the site, improved public data may become available, such as refined down-scaled climate forecasts in the region. Third, the iterative framework may allow researchers to continually adjust and refine models based on any adaptive management actions occurring within the region of the new population.

Here, we highlighted the influence of demographic structure, introgression, and dispersal on translocations; however, these
factors represent only a few potential complications genomic-informed forecasts of translocation could encounter. Other factors include habitat similarity between source and translocation sites (e.g. Wiedmann & Sargeant 2014; Batson et al. 2015), source-sink dynamics, and anthropogenic effects (Menges 2008). Accordingly, most management decisions regarding translocations take multiple factors into account, from community composition to landscape context. These different factors can interact in complex ways, making predictions of translocation success challenging and weakening or even reversing outcomes from those predicted primarily on genomics. An iterative modeling framework that tests inclusion of these additional factors will be essential for determining which factors are essential for accurate forecasts.

Conclusion
Forecasting is necessary to maximize translocation efforts for the persistence of new and established populations, and in the face of rapid, anthropogenic environmental change, the need to move organisms along the landscape will increase. Although there is a rich history of forecasting prior to translocations (Converse et al. 2013), there are few examples of incorporating genomic data in those models. Here we provided guidance on the (1) available genomic metrics; (2) limitations to genomic data; and (3) an overview of a conceptual framework and empirical examples that could integrate genomics with data on population dynamics. Iteration between models and data, including forecasting outcomes from pre-translocation data and re-evaluating models in light of post-translocation data, is the most important component of our proposed framework. Previous researchers have argued for iterative modeling during population viability analyses, where simulations are re-run as new empirical data become available (McCarthy et al. 2001). The difficulty with long-term data collection is often the lack of resources or motivation for monitoring after translocation events. The great benefit of continued data collection is the enablement of adaptive management, including learning from iterative modeling, that could improve translocation outcomes in a changing world.

Acknowledgments
This publication was made possible by the NSF Idaho EPSCoR Program and by the National Science Foundation under award number OIA-1757324. T.T.C. and A.Z. were supported by the National Science Foundation under SBE-1415297. We would also like to thank Chris Caudill for feedback on the drafts.

LITERATURE CITED
Brasier CM (2001) Rapid evolution of introduced plant pathogens via interspecific hybridization: hybridization is leading to rapid evolution of Dutch elm disease and other fungal plant pathogens. Bioscience 51:123–133
Integrating genomics and translocation with modeling


Griffith B, Scott JM, Carpenter JW, Reed C (1989) Translocation as a species conservation tool: status and strategy linked references are available on JSTOR for this article: translocation as a species conservation tool: status and strategy. Science 245:477–480


IUCN/SSC (2013) Guidelines for reintroductions and other conservation translocations. I.0. IUCN Species Survival Commission, Gland, Switzerland


Stinchcombe JR, Hoekstra HE (2008) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. Heredity 100:158–170

Coordinating Editor: Paul Nevill

Received: 28 September, 2020; First decision: 17 November, 2020; Revised: 14 February, 2021; Accepted: 15 March, 2021